UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/634,145	08/04/2003	Chew Kiat Heng	NAA 0018 PA/41049.20	5097

23368   7590   10/04/2010

DINSMORE & SHOHL LLP
FIFTH THIRD CENTER, ONE SOUTH MAIN STREET
SUITE 1300
DAYTON, OH 45402-2023

EXAMINER
WHALEY, PABLO S

ART UNIT	PAPER NUMBER
1631	

MAIL DATE	DELIVERY MODE
10/04/2010	PAPER

Please find below and/or attached an Office communication concerning this application or proceeding.

The time period for reply, if any, is set in the attached communication.

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE *3* MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.
- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1)☒ Responsive to communication(s) filed on *11 August 2010*.
2a)☐ This action is **FINAL**.   2b)☒ This action is non-final.
3)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4)☒ Claim(s) *1-15,17-26 and 28-30* is/are pending in the application.
    4a) Of the above claim(s) _____ is/are withdrawn from consideration.
5)☐ Claim(s) _____ is/are allowed.
6)☒ Claim(s) *1-15,17-26 and 28-30* is/are rejected.
7)☐ Claim(s) _____ is/are objected to.
8)☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9)☐ The specification is objected to by the Examiner.
10)☐ The drawing(s) filed on _____ is/are: a)☐ accepted or b)☐ objected to by the Examiner.
    Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).
    Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).
11)☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

12)☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).
    a)☐ All b)☐ Some * c)☐ None of:
        1.☐ Certified copies of the priority documents have been received.
        2.☐ Certified copies of the priority documents have been received in Application No. _____.
        3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).
    * See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)
1)☒ Notice of References Cited (PTO-892)
2)☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
3)☐ Information Disclosure Statement(s) (PTO/SB/08)
    Paper No(s)/Mail Date _____.
4)☐ Interview Summary (PTO-413)
    Paper No(s)/Mail Date. _____.
5)☐ Notice of Informal Patent Application
6)☐ Other: _____.

DETAILED ACTION

Request For Continued Examination

A request for continued examination under 37 CFR 1.114, including the fee set forth in

37 CFR 1.17(e), was filed in this application after final rejection.   Since this application is

eligible for continued examination under 37 CFR 1.114, and the fee set forth in 37 CFR 1.17(e)

has been timely paid, the finality of the previous Office action has been withdrawn pursuant to

37 CFR 1.114. Applicant's submission filed on 08/11/2010 has been entered.

Status of Claims

Claims 1-15, 17-26, and 28-30 are pending and under consideration.

Claims 16 and 27 are cancelled.

Withdrawn Rejections/Objections

Rejections and/or objections not reiterated from previous office actions are hereby

withdrawn in view of the amendments filed 08/11/2010. The following rejections and/or

objections are either reiterated or newly applied. They constitute the complete set presently being

applied to the instant application.

Claim rejections - 35 USC § 112, 2nd Paragraph

The following is a quotation of the second paragraph of 35 U.S.C. 112:

> The specification shall conclude with one or more claims particularly pointing out and
> distinctly claiming the subject matter which the applicant regards as his invention.

> The essential inquiry pertaining to this requirement is whether the claims set out and
> circumscribe a particular subject matter with a reasonable degree of clarity and
> particularity. Definiteness of claim language must be analyzed, not in a vacuum, but in
> light of: (A) The content of the particular application disclosure; (B) The teachings of the

prior art; and (C) The claim interpretation that would be given by one possessing the ordinary level of skill in the pertinent art at the time the invention was made.

Claims 1-15, 17-26, and 28-30 are rejected under 35 U.S.C. 112, second paragraph, as being indefinite for failing to particularly point out and distinctly claim the subject matter which applicant regards as the invention. Claims that depend directly or indirectly from claim 1, 21, and 28 are also rejected due to said dependency.

Claims 1 and 21 recite steps for optimizing the parameters of a candidate model by calculating for each of said sets, a deviate of predicted risk,..." followed by the step of "calculating a sum of weighted deviates for all said sets" (see lines 10-15). The claims previously recite a plurality of sets of data including non-genetic data, genetic data, and an indicator of disease status (lines 3-5), however the candidate model is a function of non-genetic data and parameters (see lines 8-10). In other words, the genetic data and indicators of disease status are not included in the model. Therefore, it is unclear what data is used for calculating said deviates and sum of weighted deviates. For purposes of examination, these limitations are interpreted as calculations of deviates and weighted deviates of non-genetic data, as this is the only type of data required by the model.

Claims 1 and 21 recite a step for "determining the weights used to weight said deviates with a constraint that said weights associated with sets of said data having like genetic data are the same" (lines 17-onward). It is unclear what limitation of the claimed method is intended. For example, one could interpret this to be determining weights with the constraint that the weights are the same. Another interpretation could be determining weights under the constraint that the weights are restricted to a subpopulation of genetic data, wherein the genetic data is restricted to

be the same. As a result, it is unclear in what way the condition "with a constraint wherein genetic data is the same" is incorporated into the calculation of weights. Furthermore, a review of the specification [0062-0074] shows weights determined by grouping, partitioning data into groups and using genetic markers, and tree-pruning procedures, all of which are inconsistent with determining weights based on a condition, as recited in the claims. For purposes of examination, this limitation is interpreted as determining weights under the constraint that the weights associated with genetic data are the same.

Claim 28 recites calculating "a sum of weighted deviates, each deviate weighted by a weight reflecting genetic data associated with that member for whom that deviate is calculated" (lines 11-15). It is unclear what is meant by a weight "reflecting genetic data." Is the weight determined from the genetic data; e.g. via some mathematical relationship? Is this an intended use of the weight? Clarification is requested.

NEW GROUND OF REJECTIONS

Claim Rejections - 35 USC § 103

The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

(a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the manner in which the invention was made.

This application currently names joint inventors. In considering patentability of the claims under 35 U.S.C. 103(a), the examiner presumes that the subject matter of the various claims was commonly owned at the time any inventions covered therein were made absent any evidence to the contrary. Applicant is advised of the obligation under 37 CFR 1.56 to point out the inventor and invention dates of each claim that was not commonly owned at the time a later invention was made in order for the examiner to consider the applicability of 35 U.S.C. 103(c) and potential 35 U.S.C. 102(e), (f) or (g) prior art under 35 U.S.C. 103(a).

The factual inquiries set forth in Graham v. John Deere Co., 383 U.S. 1, 148 USPQ 459 (1966), that are applied for establishing a background for determining obviousness under 35 U.S.C. 103(a) are summarized as follows:
1.   Determining the scope and contents of the prior art.
2.   Ascertaining the differences between the prior art and the claims at issue.
3.   Resolving the level of ordinary skill in the pertinent art.
4.   Considering objective evidence present in the application indicating obviousness or nonobviousness.

Claims 1-3, 9, 13-15, 17-21, and 28-30 are rejected under 35 U.S.C. 103(a) as being made obvious by Walter et al. (American Journal of Epidemiology, 1978, 108(5): 341-6), in view of Pharaoh et al. (Nature Genetics; May 2002, 31: 33-36), Shattuck-Eidens et al. (JAMA, 1997, 278 (15): 1242-1250), and in view of Pfeffermann et al. (International Statistical Review, 1993, 61, 2:317-337).

The claims have been amended and are now drawn to a computer-implemented method of determining a statistical model for predicting disease risk for a member of a population. Critical limitations of claims 1 and 21 include collecting a plurality of data sets at a computing device. Each data set is associated with one member of a population and includes non-genetic data, genetic data that is indicative of the presence or absence of a genetic marker in said one member, and data that indicates disease status. A candidate statistical model for calculating disease risk as a function of non-genetic data is stored at a computing device. The model

parameters are optimized by at least one computing device by fitting. The fitting is based on calculating a deviate of a predicted risk from an indicator of disease status for each set by using the candidate model and non-genetic data; calculating a sum of weighted deviates for all sets, where the weights are associated with the set for which each deviate has been calculated; and determining weights used to weight the deviates with a constraint such that sets with the same genetic data have the same weights. Additionally, the optimum parameters are obtained by minimizing the sum of weighted deviates used with the candidate model for calculating disease risk. Claims 20 and 21 are directed to a program and system for performing the above method steps.

Walter teaches methods of determining models for predicting disease risk for members of a population. In particular, Walter teaches a plurality of statistical models for predicting risk, for example based on two causal risk factors A and B and a collective risk C, to which a population is exposed; see pages 341-343, Examples 1 and 2. For each exposure group, the risk of one or more adverse events (i.e. disease) is defined by probability functions P(A), P(B), and P(C); see page 342, Col. 1. Thus, probability functions can be obtained for any additive combination of the known risk factors A, B, and C. A log-linear model is also presented; see p.345. The modeled data is multidimensional, where one of the dimensions corresponds to disease status (present or absent) and the other dimensions to the presence or absence of various risk factors; see p.344, Col. 2. The presence and absence of all variables is denoted by 1 and 0; see p. 345, Col. 1. Goodness-of-fit tests statistics (i.e. sum of weighted deviates) are used to assess model accuracy; see p.344, Col. 2. The appropriate model is determined by its ability fit data, and criteria for defining a good fit vary according to methods used for estimating parameters; see p.346,

Summary. In particular, iterative maximum-likelihood estimates and weighted least squares criterion may easily be performed for optimizing model parameters by fitting; see p.346. The candidate model can be represented as an exponential function and can represent disease at a given time; see p.342, Col. 1.

Walter does not specifically teach data that reflect the use of non-genetic data and genetic data, as in claims 1 and 21, which are interpreted as non-genetic and genetic risk factors in view of the specification [0031].

Walter does not teach optimizing model parameters based on calculating a sum of weighted deviates for all data sets under the constraint that the weights are restricted to a subpopulation of genetic data that is the same, as in claims 1 and 21.

Walter does not teach optimizing model parameters using an adjustment factor that is a function of a ratio of the number of members in a population who share the same characteristics over the total number of members in a population, as in claims 13 and 14.

Methods for predicting risk using non-genetic and genetic risk factor data would have been known in the art. Pharaoh teaches a statistical model for predicting cancer risk. In particular, Pharaoh shows collecting genetic and non-genetic risk factor data (i.e. genetic and non-genetic data); see Abstract, p.35, Col. 1, and Table 1. Polygenic models for predicting risk in a population are fitted to the population data; see pages 33-34. The models determine the risk of breast cancer conferred by single genes; see Table 1 and p. 35, Col. 2, which shows genetic data indicative of the presence or absence of genetic markers.

Shattuck-Eidens teaches a method for predicting the presence of harmful genetic mutations in individuals [see p.1243, Col. 3, p.1244, Col. 1 and 2, and p.1246, Col. 3]. In

particular, non-genetic data and indicators of disease status is collected [p.1243-1244 and p.1246, Col. 2 and 3]. Data is obtained for genetic markers indicative of the early onset of breast cancer [Tables 2, 3, and 4, and p.1249, Col. 2], which shows genetic markers indicative of the presence of disease. The probability of carrying a genetic mutation is calculated using a regression model. that is a function of non-genetic data and parameters; see p. 1243, Col. 3]. Optimization is performed by fitting model parameters based on additive scaled deviance calculations; see p. 1243, Col. 3 and p.1244, Col. 3, which is interpreted as fitting of model parameters by calculating a sum of weighted deviance of a predicted risk using non-genetic data. Additionally, populations are divided into groups based on their genetic data, wherein at least one of the groups has the same genetic mutation; see p. 1244, Col. 3, which is interpreted as a group having the same genetic data. Furthermore, groups with the same disorder are assigned a similar integer value [p.1246, Col. 2 and 3], which reasonably suggest a constraint such that sets with the same genetic data have the same weights. The above analyses are performed using a software package; see p.1244, Col. 2, which shows collecting data by a computing device, and an article of manufacture and system for performing the above method steps.

Pfeffermann teaches methods for using sampling weights in statistical models. In particular, Pfeffermann discusses the use of sampling weights for estimating model parameters by fitting; see p.318, Section 4.2, and p.327. In one case, the weight represented by a step function that equates to 0 or 1 based on a particular constraint where the data is great than or equal to 0; see p.327, which shows weights based on a constraint. Samples can also be weighted using a count estimate of the form Ni/Nj, which shows an adjustment factor that is a function of a ratio of the number of members in a population who share the same characteristics over the

total number of members in a population The motivation would have been to correct for disproportionality of the sample in the target population [Introduction].

It would have been obvious to someone of ordinary skill in the art at the time of the instant invention to have used non-genetic and genetic risk factor data in the disease risk in the model of Walter, with a reasonable expectation of success, since Walter teaches a generic risk model that can use any type of causal risk factor and binomial data, as set forth above, and since Pharaoh and Shattuck-Eidens teach risk models that use non-genetic and genetic risk factor data, as set forth above. The motivation to use a combination of known genotype and non-genotype risk factors would have been to provide risk discrimination that has practical value for health care, as suggested by Pharaoh; see page 35, Col.2.

It would have been obvious to someone of ordinary skill in the art at the time of the instant invention to have optimized model parameters based on calculating a sum of weighted deviates for all data sets under the constraint that the weights are restricted to a subpopulation of genetic data that is the same in the disease risk in the method made obvious by Walter, Pharaoh, and Shattuck-Eidens with a reasonable expectation of success, since Shattuck-Eidens and Walters shows model optimization based on sums of weighted deviates, as set forth above, since Shattuck-Eidens suggests assigning the same weights to groups with similar disorders [p.1246, Col. 2 and 3], which suggests a constraint such that sets with the same genetic data have the same weights, and since Pfeffermann shows that weighted samples can be predictably used for estimating model parameters by fitting; see p.318, Section 4.2, and p.327. The motivation would have been to correct for disproportionality of the sample in the target population [Pfeffermann, Introduction].

It would have been obvious to someone of ordinary skill in the art at the time of the instant invention to have optimized model parameters using an adjustment factor that is a function of a ratio of the number of members in a population who share the same characteristics over the total number of members in a population, in the method made obvious by Walter, Shattuck-Eidens, Pharaoh, and Pfeffermann, with a reasonable expectation of success, since Pfeffermann also shows weighted data using a count estimate of the form Ni/Nj, as set forth above, which is interpreted as an adjustment factor that is a function of a ratio of the number of members in a population who share the same characteristics over the total number of members in a population, and in view of the rationale for a *prima facie* case of obviousness provided by the Supreme Court in *KSR International Co. v. Teleflex Inc.*, 550 U.S. 398 (2007). See MPEP 2143. In this case, the rationale would have been to improve model fitting based on variations of known design parameters, such as using population ratios to decrease the variance of the estimators, as suggested by Pfeffermann; see p.329, since these variations are predictable to one of ordinary skill in the art. For these reasons, the instant claims do not recite any new element or new function or unpredictable result.

Claims 4, 5, and 23-26 are rejected under 35 U.S.C. 103(a) as being made obvious by Walter et al. (American Journal of Epidemiology, 1978, 108(5): 341-6), in view of Pharaoh et al. (Nature Genetics; May 2002, 31: 33-36), Shattuck-Eidens et al. (JAMA, 1997, 278 (15): 1242-1250), and Pfeffermann et al. (International Statistical Review, 1993, 61, 2:317-337), as applied to claims 1-3, 9, 13-15, 17-21, and 28-30, and further in view of Nelson et al. (J Clin Epidemiol, 1998, Vol. 51, No. 3, pp. 199–209), and Marshall et al. (Statistics in Medicine; 1986;5:517-526).

Walter, Shattuck-Eidens, Pharaoh, and Pfeffermann make obvious a method and program

for determining a statistical model for predicting disease risk, as set forth above.

Walter, Shattuck-Eidens, Pharaoh, and Pfeffermann do not teach grouping collected data

such that all sets of data have like genetic data, one of said group being a reference group, and

determining a group weight for each group, wherein the group weight is between 0 and 1, as in

claims 4 and 5.

Walter, Shattuck-Eidens, Pharaoh, and Pfeffermann do not teach dividing sets of data

into two or more groups depending on data indicative of non-genetic factors, determining if a

criterion is met after dividing, and regrouping sets of data back into one group when criterion is

not met, as in claim 23-26.

Nelson teaches method of recursive partitioning for the identification of disease risk

groups; Abstract. Individual groups of data are split by recursive partitioning in order to identify

variables that minimize the variance between the subsets; see p.201-202 and Fig. 1.

Classification tree models are constructed and an optimum sized model is selected using cross-

validation; see p.201, Col. 1 and Col. 2. The method requires fitting dummy variables with

conditional logistic regression to estimate odds ratio's (i.e. group weights) for subsets in a

classification tree; see p.204, Col. 2. Each subset that is partitioned in classification tree is

assigned an index value based on a weighted average; see Appendix A, which is interpreted as an

alternative teaching for group weights between 0 and 1; see p.207, Col. 2. The partitioning of

data is based on all of different risk factor groups involved in the study; e.g. case and control

data, see p.202 and Fig. 1. An optimized tree is obtained as a result of partitioning, pruning, and

cross-validation, and that classification tree risk subgroup can also be modeled via conditional

regression analysis, which inherently include optimization of model parameters by minimizing of

target functions; see page 204, Col. 1, and Col. 2. The motivation would have been to allow

effect estimates to be adjusted for matching variables; p.204, Col. 1.

Marshall teaches methods of recursive partitioning of data. In particular, Marshall teaches

combining partitions based on search criteria; see p.522, which is interpreted as regrouping data.

The benefit of this method is to obtain optimal partitions. Group data is assigned to binary

values; e.g. 0 and 1; see page 518-519.

It would have been obvious to someone of ordinary skill in the art at the time of the

instant invention to have grouped data such that all sets have the same genetic data and one of

said groups is a reference group, in the method made obvious by Walter, Shattuck-Eidens,

Pharaoh, and Pfeffermann, with a reasonable expectation of success, in view of the teachings of

Pharaoh, who shows groups that are susceptible to different genetic factors have different

probabilities of risk; see page 34, Col. 1, and since Nelson shows the probabilities can be

predictably determined for case and control groups, see p.201, Col. 2. The motivation would

have been to minimize the variation in predicted risk, as suggested by Pharaoh, p.34, Col. 1.

It would have been obvious to someone of ordinary skill in the art at the time of the

instant invention to have determined group weights, wherein weights are between 0 and 1, in the

method made obvious by Walter, Shattuck-Eidens, Pharaoh, and Pfeffermann, with a reasonable

expectation of success, since Nelson and Marshall both show methods for partitioning data into

groups assigning binary values (i.e. weights) to the groups, as set forth above. The motivation

would have been to uncover interactions between variables that may be overlooked in the

traditional application of logistic regression to case-control data, as suggested by Nelson; Summary and p. 208, Col. 1.

It would have been obvious to someone of ordinary skill in the art at the time of the instant invention to have optimized group weights by minimizing a target function that is dependent on a plurality of residuals, in the method made obvious by Walter, Shattuck-Eidens, Pharaoh, and Pfeffermann ,with a reasonable expectation of success, since Nelson shows fitting variables with conditional logistic regression to estimate odds ratio's (i.e. group weights) with predictable results, as set forth above, which suggests optimizing group weights by minimizing a target function. The motivation would have been to improve results through standard methods of model optimization, as suggested by Nelson; see p. 201.

It would have been obvious to someone of ordinary skill in the art at the time of the instant invention to have divided sets of data into two or more groups depending on data indicative of non-genetic factors, and determined if a criterion is met after dividing, in the method made obvious by Walter, Shattuck-Eidens, Pharaoh, and Pfeffermann, with a reasonable expectation of success, since Nelson shows dividing non-genetic data into groups (i.e. case and control) based on a numerical criterion, and repeats the partitioning process for all subgroups after the division process; see p.201, Col. 2. The motivation would have been to perform routine optimization for identifying the variable that gives the best separation of case and control groups, as suggested by Nelson; p.208, Col. 1.

It would have been obvious to someone of ordinary skill in the art at the time of the instant invention to have regrouped sets of data back into one group when criterion is not met, in the method made obvious by Walter, Shattuck-Eidens, Pharaoh, and Pfeffermann, with a

reasonable expectation of success, since Marshall shows recursive partitioning of data by combining partitions based on search criteria; see p.522, which is interpreted as regrouping data. The motivation would have been to obtain optimal partitions in statistical models, as suggested by Marshall, above.

Claims 6, 7, 8, 9, 10, and 11 are rejected under 35 U.S.C. 103(a) as being made obvious by Walter et al. (American Journal of Epidemiology, 1978, 108(5): 341-6), in view of Pharaoh et al. (Nature Genetics; May 2002, 31: 33-36), in view of Shattuck-Eidens et al. (JAMA, 1997, 278 (15): 1242-1250), in view of Pfeffermann et al. (International Statistical Review, 1993, 61, 2:317-337), in view of Nelson et al. (J Clin Epidemiol, 1998, Vol. 51, No. 3, pp. 199–209), and in view of Marshall et al. (Statistics in Medicine; 1986;5:517-526), as applied to claims 1-3, 9, 13-15, 17-21, and 28-30, and further in view of Parzen et al. (Biometrics, 1999, 55, 580-584).

Walter, Shattuck-Eidens, Pharaoh, Pfeffermann, Nelson, and Marshall make obvious a method and program for determining a statistical model for predicting disease risk, as set forth above.

Walter, Shattuck-Eidens, Pharaoh, Pfeffermann, Nelson, and Marshall do not teach minimizing a target function dependent on residuals and weights, as in claims 6, 7, and 8.

Walter, Shattuck-Eidens, Pharaoh, Pfeffermann, Nelson, and Marshall do not teach data indicative of time or using a Cox hazard model, as in claims 9 and 10.

Parzen teaches a Cox hazard regression model for calculating disease risk in a subject based on non-genetic data as a function of time [Section 2, Table 1, Table 2, and p.581, Col. 2].

Parzen calculates partial risk estimates using a Cox likelihood score vector based on a sum of weighted averages and a binary variable between 1 and 0 (i.e. weight) [p.581, Col. 2, Equation 2], which is interpreted as a minimizing a target function. Model parameters are optimized by data fitting [Section 3], and using a residual equation for calculating goodness of fit [p.582, Col. 2].

It would have been obvious to someone of ordinary skill in the art at the time of the instant invention to have minimized a target function dependent on residuals and weights, as taught by Cleveland, above, in the method made obvious by Walter, Shattuck-Eidens, Pharaoh, Pfeffermann, Nelson, and Marshall, with a reasonable expectation of success, since Parzen also teaches optimization of regression model parameters, as set forth above. The motivation would have been to improve model accuracy using an alternative scoring statistic, as suggested by Parzen; p.582, Col. 2.

It would have been obvious for one of ordinary skill in the art at the time of the instant invention to have provided a predictable variation of the type regression model used, such as the Cox hazard regression model taught by Parzen, in the method made obvious by Walter, Shattuck-Eidens, Pharaoh, Pfeffermann, Nelson, and Marshall, with a reasonable expectation of success, in view of the rationale for a *prima facie* case of obviousness provided by the Supreme Court in *KSR International Co. v. Teleflex Inc.*, 550 U.S. 398 (2007). See MPEP 2143. In this case, the rationale would have been the simple substitution of one known method of predicting risk for another, such as a Cox regression model, since these variations are predictable to one of ordinary skill in the art. For these reasons, the instant claims do not recite any new element or new function or unpredictable result.

Claims 12 and 22 are rejected under 35 U.S.C. 103(a) as being made obvious by Walter et al. (American Journal of Epidemiology, 1978, 108(5): 341-6), in view of Pharaoh et al. (Nature Genetics; May 2002, 31: 33-36), Shattuck-Eidens et al. (JAMA, 1997, 278 (15): 1242-1250), and Pfeffermann et al. (International Statistical Review, 1993, 61, 2:317-337), as applied to claims 1-3, 9-11, 13-15, 17-21, and 28-30, and further in view of Raghunathan et al. (Survey Methodology, 2001, 27(1):85-95).

Walter, Shattuck-Eidens, Pharaoh, Pfeffermann, Nelson, and Marshall make obvious a method and program for determining a statistical model for predicting disease risk, as set forth above.

Walter, Shattuck-Eidens, Pharaoh, Pfeffermann, Nelson, and Marshall do not teach imputing missing data, as in claims 12 and 22.

Raghunathan teaches a method for imputing missing data into a statistical model; see Abstract, Section 3, p.89, Col. 1. Regression is performed for a data set that has missing data under a constraint that a subpopulation is the same; see p.89, Col. 1. The logistic regression model is then fitted to each imputed data set to obtain maximum likelihood estimates of the regression coefficients and asymptotic covariance matrices; see p.89, Col. 2. The method uses "multiply imputed estimates" of the regression coefficients calculated as a ratio; see p.89, Col. 2, which are interpreted as determining adjustment factors under a constraint.

It would have been obvious to someone of ordinary skill in the art at the time of the instant invention to have imputed missing data into a model, as taught by Raghunathan, above, in

the method made obvious by Walter, Shattuck-Eidens, Pharaoh, Pfeffermann, Nelson, and Marshall, with a reasonable expectation of success, since Shattuck-Eidens shows grouping of data; see p. 1244, Col. 3, and suggests the existence of incomplete data sets; Table 5 and p.1248, Col. 2. The motivation would have been to improve model accuracy when data is missing at random, as suggested by Raghunathan [Abstract].

Response to Arguments

Applicant's arguments filed 8/11/2010 with respect to the rejections of claims 1-6, 9-11, 13-15, 17-21, and 28-30 under 35 USC 103 in view of Parzen, Shattuck-Eidens, and Cleveland have been fully considered. Applicant's arguments that Shattuck-Eidens does not teach calculating a sum of weighted deviates, where each deviate is weighted by a weight reflecting genetic data are persuasive. Therefore, the rejection has been withdrawn. However, upon further consideration, a new ground of rejections has been applied, as set forth above.

Conclusion

Any inquiry concerning this communication or earlier communications from the examiner should be directed to Pablo Whaley whose telephone number is (571)272-4425. The examiner can normally be reached between 12pm-8pm.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Marjorie Moran can be reached at 571-272-0720. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see http://pair-direct.uspto.gov. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free).

Pablo S. Whaley
Patent Examiner
Art Unit 1631

/PW/

/Marjorie Moran/
Supervisory Patent Examiner, Art Unit 1631